Meteorology and Atmospheric Physics (2021) 133:943–959
https://doi.org/10.1007/s00703-021-00791-4

ORIGINAL PAPER

Short-term air temperature prediction by adaptive neuro-fuzzy inference system (ANFIS) and long short-term memory (LSTM) network

Aliihsan Sekertekin[1] · Mehmet Bilgili[2] · Niyazi Arslan[1] · Alper Yildirim[3] · Kerimcan Celebi[2] · Arif Ozbek[2]

Received: 18 June 2020 / Accepted: 7 March 2021 / Published online: 20 March 2021

Abstract
Air Temperature (AT) is a crucial parameter for many disciplines such as hydrology, irrigation, ecology and agriculture. In this respect, accurate AT prediction is required for applications related to agricultural operations, energy generation, traveling, human and recreational activities. In this study, four different machine learning approaches such as Adaptive Neuro-Fuzzy Inference System (ANFIS) with Fuzzy C-Means (FCM), ANFIS with Subtractive Clustering (SC) and ANFIS with Grid Partition (GP) and Long Short-Term Memory (LSTM) neural network were used to make one-hour ahead and one-day ahead short-term AT predictions. Concerning the test site, the measured AT data were obtained from a solar power plant installed in the city of Tarsus, Turkey. Correlation coefficient (*R*), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) were used as quality metrics for prediction. Predicted values of the LSTM, ANFIS-FCM, ANFIS-SC and ANFIS-GP models were compared with the observed values by evaluating their prediction errors. According to the hourly AT prediction, the RMSE values in the testing process were found to be 0.644 (°C), 0.721 (°C), 0.722 (°C) and 0.830 (°C) for the LSTM, ANFIS-FCM, ANFIS-SC and ANFIS-GP models, respectively. On the other hand, the RMSE values of the corresponding methods for daily AT prediction were obtained as 1.360 (°C), 1.366 (°C), 1.405 (°C) and 1.905 (°C), respectively. The comparison of hourly and daily prediction results revealed that the LSTM neural network provided the highest accuracy results in both one-hour ahead and one-day ahead short-term AT predictions, and mainly presented higher performance than all ANFIS models.

1 Introduction

Global warming, related with an increase of average global air temperature (AT) near the Earth's surface above the continents and sea surface temperature above oceans in comparison to the period of preindustrial period 1850–1900, is very probably a consequence of human activities. Because of simplicity, this hybrid global average air and ocean temperature is called as "average global surface temperature" and may result in climate change in the world (Christensen et al. 2007). In recent years due to this climate change, extreme natural events such as heat waves, severe colds, heavy snow and rain, and drought have been on rising in all regions of the world, causing environmental and health problems, material damage, injuries, and deaths. AT value that can change at any time is an important parameter of the air (Zahroh et al. 2019). In agricultural processes, crop growth and yield are significantly affected by the AT variations (Meshram et al. 2020). AT can also be used for simulation of food safety and crop production under changing climatic conditions (Hoogenboom 2000). Furthermore, there is a strong correlation between temporal AT changes and energy consumption, agricultural activity, livestock farming, water consumption, and atmospheric events (Nag et al. 2009). Considering the productivity in agriculture, AT prediction can play a significant role in planning agricultural activities, such as planting and harvesting dates of the crop, irrigation, pruning, and frost preventing.

Predictive analytics is a technique of logical analysis of data aimed at making predictions about future outcomes based on historical data. Predictive analytics covers a wide

Responsible Editor: Maja Telisman Prtenjak.

✉ Mehmet Bilgili
mbilgili@cu.edu.tr

[1] Department of Geomatics Engineering, Cukurova University, 01950 Adana, Turkey

[2] Department of Mechanical Engineering, Cukurova University, 01950 Adana, Turkey

[3] Department of Machinery and Metal Technology, Osmaniye Korkut Ata University, 80000 Osmaniye, Turkey

range of approaches, including big data analysis, machine learning, statistical modeling, and assorted mathematical processes. Future predictions can be more reliable and accurate than previous tools. The benefits of predictive analytics include swifter, increased competitiveness, smarter decision-making, greater agility, and improved risk management. This method helps people in various fields such as finance, hospitality, healthcare, automotive, pharmaceuticals, energy, manufacturing, and aerospace (Mike Capone 2020).

Weather data prediction, which has drawn the interest of the scientists, has a significant impact on human life (Kumar et al. 2020). Thus, many researches have been carried out to predict the meteorological parameters, and numerous publications have emerged on the subject (Salman et al. 2018; Bilgili 2010; Chaudhuri and Middey 2011; Liu et al. 2020). However, this process can be complicated and challenging for the researchers concerning the accuracy of the results (Radhika and Shashi 2009). Accurate results can be obtained using artificial intelligence and statistical analysis techniques when predicting AT for short-term time interval (hourly, daily). The variation between the observed and the predicted temperature or the error increases for the long-term prediction (Salman et al. 2018). Li et al. (2020) developed a number of machine learning and statistical techniques for establishing climate predictions, and compared their performance for downscaling long-term daily air temperature. Recently, various prediction models have been proposed to improve AT and Land Surface Temperature (LST) prediction performance. These models fall into six categories: the Spatio-Temporal Correlation (STC) model, the deep learning model, the machine learning model, the combined model, the statistical model and the physical model (Liu et al. 2018).

Weather prediction is affected by high dimensionality, synergies on many distinct temporal and spatial scales, and uncontrolled dynamics. This effect complicates many problems in the field. Also, despite being state of the art, digital weather forecast (NWP) can be not sufficient for many applications. Therefore, it is quite appropriate to use emerging new approaches such as artificial neural networks (ANN) to overcome these problems (Scher 2020). ANN mimic the human brain in processing input signals and transform them into output signals. It provides a powerful modeling algorithm that allows for non-linearity between feature variables and output signals. ANN is a kind of non-parametric modeling technique, which is suitable for the complex phenomenon that investigators do not know underlying functions (Zhang 2016). These features make the ANN techniques very appealing in application domains for solving highly non-linear phenomena. Deep models for multivariate time-series forecasting often use Temporal Convolutional Networks (TCN) and Recurrent Neural Networks (RNN). A variant of RNN called Long Short-Term Memory (LSTM) has attached significant attention due to its high performance (Hewage et al. 2020). To better understand ANN, it can be compared to the conventional regression model. The primary advantage of ANN over the conventional regression analysis is that they are capable of processing vast amounts of data and making predictions without having to know the statistical background (Bandara et al. 2011).

Considering the above-mentioned prediction studies, Cobaner et al. (2014) applied ANNs, Multiple Linear Regression (MLR) and ANFIS models to predict monthly minimum, maximum and average ATs for any geographic region of Turkey. In this study, it was stated that the best estimation result was acquired from ANFIS. LSTM model was developed for the prediction of AT in Bandung, Indonesia (Zahroh et al. 2019). Compared to the LSTM model used with long-term data, short-term temperature prediction was largely accurate. Zhang et al. (2018) used a new hybrid method called EEMD-LSTM; in which Ensemble Empirical Mode Composition (EEMD) and LSTM network was utilized together to recommended daily LST data, to succeed in the modeling difficulty, and to increase the forecasting accuracy. They compared their findings with the other models used in the study; coming to a conclusion that the most accurate results were obtained from the EEMD-LSTM model. Zhang et al. (2014) used the prediction model, which considers the physical relationship between AT, heat convection, clouds and ground heat radiation, and direct solar radiation to provide temperature control in buildings. This model provided less than 1 K error and can be included in the Heating, Ventilation and Air Conditioning (HVAC) system.

Arslan and Sekertekin (2019) reconstructed Moderate Resolution Imaging Spectroradiometer (MODIS) LST images with LSTM network model using consecutive LST images taken from a selected area in Ceyhan/Adana region. Misra et al. (2018) used used Recurrent Neural Network (RNN) with LSTM to find out the spatio-temporal dependencies in local rainfall and to assess the hydrological influences of global climate change on a regional scale. Venkadesh et al. (2013) proposed weather prediction model based on the Genetic Algorithm (GA) that gives better accuracy results than the other prediction models. Radhika and Shashi (2009) developed the Support Vector Machines (SVMs) method for atmospheric temperature prediction. In that study, non-linear regression model was used to train the SVM. Sekula et al. (2019) created the Numerical Weather Prediction (NWP) and Aire Limitée Adaptation Dynamique Développement International High-Resolution Limited Area Model (ALADIN-HIRLAM) systems to predict the AT. These systems were applied to the Western Carpathian Mountains in Poland, using two different vertical and horizontal resolutions. From the comparison of the results, there was no statistically significant difference between the model results and the observations based on Root Mean Square Error (RMSE). Azad et al. (2019) represented the

importance of the minimum, average and maximum AT prediction for the water resources and their management using ANFIS model in Iran. Benmouiza and Cheknane (2019) suggested ANFIS models with FCM, SC, and GP structures to predict hourly solar radiation.

Support Vector Regression (SVR), one of the supervised learning methods, has also been implemented in prediction-based studies. Chevalier et al. (2011) developed the SVR method to make short-term air temperature prediction using meteorological data from Republic of Georgia. They suggested that SVR model can be utilized as an alternative to ANNs model in the air temperature assessment. Bilgili and Sahin (2010) applied ANNs to develop an assessment of long-term monthly precipitation and temperature from the neighboring measuring stations data in Turkey. Geographical variables (longitude, latitude, altitude) and time were considered as input data for this approach. The researchers stated that with the use of ANN method, the long-term monthly precipitation amount and temperature value of any region could be assessed without measurement depending on geographical variables. Celebi et al. (2017) were interested in the use of the ANN method to model the operation, performance, emissions and even developed the operational maps for the diesel engines. Ramesh and Anitha (2014) proposed a prediction system based on the seven-day maximum and minimum surface air temperature for the station in Chennai, India using Multivariate Adaptive Regression Spline (MARS) method. The accuracy results were compared with the SVR model. The excellent success of artificial intelligence in predictions has also attracted the attention of other fields, as well. Peng et al. (2018) utilized the LSTM with differential evolution (DE-LSTM) algorithm model in electricity price estimation.

Recently, some researches have shown that the hybrid prediction mechanisms outperform single models. To improve the performance of predictions, more sophisticated prediction approaches than traditional machine learning methods such as SVM and ANN are required (Chen et al. 2018). Qin et al. (2019) used LSTM network hybrid model for wind turbine signal prediction. The efficiency of the method has been verified by the real wind field data. The method yielded only 5% prediction error of less than 1 m for very short-term prediction periods. Han et al. (2019) proposed LSTM network and copula function to predict medium and long-term wind and photovoltaic power generation. Zhou et al. (2019) carried out the remaining useful life estimation of the super capacitor with a LSTM RNN. With this method, the effect of temperature and tension on the aging tendency of the super-capacitors was analyzed. Ma et al. (2015) applied the LSTM-RNN model for capturing the non-linear traffic dynamics. LSTM RNN yielded the best performance during the research regarding stability and accuracy.

With regard to the above-mentioned studies, in these days, the deep learning methods have been performed excellently in high-accuracy prediction studies, and therefore have received intense attention of researchers. In the deep learning process, RNNs are well suited to solve sequence problems such as time-series data. The LSTM network algorithm offers the special form and development of the RNN. However, it has a much better universality than that of traditional RNN. The LSTM has been successfully utilized in different applications, especially in time sequence problems. Since AT variations significantly affect the crop growth and yield in agricultural processes, in this study four data-driven methods, the ANFIS-FCM, ANFIS-SC and ANFIS-GP and LSTM are proposed for one-hour ahead and one-day ahead short-term AT predictions. The one-hour intervals and daily AT data series obtained from the Pilye Solar Power Plant in the Mediterranean Region of Turkey were used as a sample dataset. Three statistical error criteria (MAE, RMSE and R) were used to evaluate the performance of the methods.

The main original contributions to the literature of this study can be primarily expressed in four aspects: (i) An LSTM neural network, based on a deep learning approach, is proposed to forecast the one-hour ahead and one-day ahead air temperature predictions. Regarding the published forecasting models on AT in the literature, there are many studies based on the approaches such as machine learning models, time-series models and hybrid models. However, the number of studies on the forecasting of AT using the LSTM neural network is limited. (ii) The proposed LSTM neural network is demonstrated to be extremely appropriate in the forecasting of complex time series with seasonality. (iii) To show the superiority of the approach used in this study, the results obtained from the LSTM neural network are compared with the findings of the ANFIS models based on the same data sets used in this study. (iv) In addition, the one-hour ahead AT prediction results are compared with the one-day ahead ones to show which data set is more feasible with the corresponding methods.

2 Methodology

2.1 Adaptive neuro-fuzzy inference system (ANFIS)

ANFIS, a global predictor, has the capability for obtaining high accuracies with several kinds of actual and repeated phenomena. Basically, ANFIS, which is equipped with neural learning capabilities, is expressed as Sugeno-type fuzzy systems' network statement. It creates a cluster of fuzzy if–then rules via suitable Membership Functions (MFs) from the pairs of input and output data using a neural network learning algorithm. This Fuzzy Inference

System (FIS) development strategy that uses the adaptive neural network framework is called ANFIS (Abyaneh et al. 2011).

Jang (1993) presented the learning structure and the architecture of ANFIS. According to Jang (1993), the ANFIS model functionally has the same structure as the Takagi–Sugeno-type extraction model. An ANFIS model is a combination of two statistical systems, namely, FIS and ANN. ANFIS can model the dynamics of difficult systems in a best way. At first, the dataset is trained by ANFIS model in a similar manner as ANN. The trained system is then operated as a FIS. ANFIS combines both ANN and FIS bases taking into account the advantages of both systems into a single system (Karakuş et al. 2017).

The neuro-fuzzy model in Tabari et al. (2012) is given in Fig. 1 representing a typical ANFIS architectural structure. It has a total of five layers as a fuzzy system based on multi-layered neural network. In the structure of this network, input and output nodes refer to input conditions states and output response, respectively, as well as nodes in hidden layers that act as rules and MFs. Thus, this phenomenon removes the drawback of a traditional feed forward multi-layered network that is challenging for a supervisor to figure out or change. In this structure, a circle shape stands for a fixed node while a square shape represents an adaptive node. Simply, in Fig. 1, x and y are two inputs, and z refers to one the output. In this study, we have one input and one output datasets that are both AT time series. Thanks to its built-in optimal and adaptable techniques, computational efficiency and high interpretability, the Sugeno model is the most widely applied among many fuzzy models in the literature. Two fuzzy if–then rules depending on the first-order Sugeno fuzzy model can be given by:

$$\text{Rule 1}: \text{if } x \text{ is } A_1 \text{ and } y \text{ is } B_1, \text{ then } z_1 = p_1 x + q_1 y + r_1, \quad (1)$$

$$\text{Rule 2}: \text{if } x \text{ is } A_2 \text{ and } y \text{ is } B_2, \text{ then } z_2 = p_2 x + q_2 y + r_2, \quad (2)$$

where B_i and A_i express the fuzzy clusters in the first one, while r_i, q_i and p_i define the architectural parameters in the training process. Figure 1 illustrates that five layers, detailed below, constitute the ANFIS structure (Tabari et al. 2012):

Layer 1: This layer defines an input variable for each appropriate fuzzy set. The function by which membership degrees are created using MFs is indicated by the nodes in this layer. The node function of a node i is defined as follows:

$$O_i^1 = \mu_{A_i}(x), \ i = 1, 2, \quad (3)$$

$$O_i^1 = \mu_{B_{i-2}}(y), \ i = 3, 4, \quad (4)$$

where μ_{A_i} and μ_{B_i} are the MFs. In this layer, the parameters are defined as preliminary parameters.

Layer 2: In the nodes of Layer 2, incoming signals are duplicated and the product out is sent. In this layer, all of the nodes computes the firing strength of a rule by multiplying the incoming signals.

$$O_i^2 = w_i = \mu_{A_i}(x)\mu_{B_i}(y), \ i = 1, 2. \quad (5)$$

Layer 3: The ith node in Layer 3 computes the ratio of the ith rule's firing strength to the sum of all rules firing strengths of all rules:

$$O_i^3 = \overline{w}_i = \frac{w_i}{w_1 + w_2}, \ i = 1, 2, \quad (6)$$

where $\overline{w}_i$ means the normalized firing strengths.

Layer 4: Node i in Layer 4 determines how the ith rule will contribute to the model output, via the node function in Eq. (7)

$$O_i^4 = \overline{w}_i z_i = \overline{w}_i (p_i x + q_i y + r_i), \ i = 1, 2, \quad (7)$$

where $\{r_i, q_i \text{ and } p_i\}$ is the parameter set, and $\overline{w}_i$ is the output of the layer 3. In this layer, the parameters are called the consequent parameters.

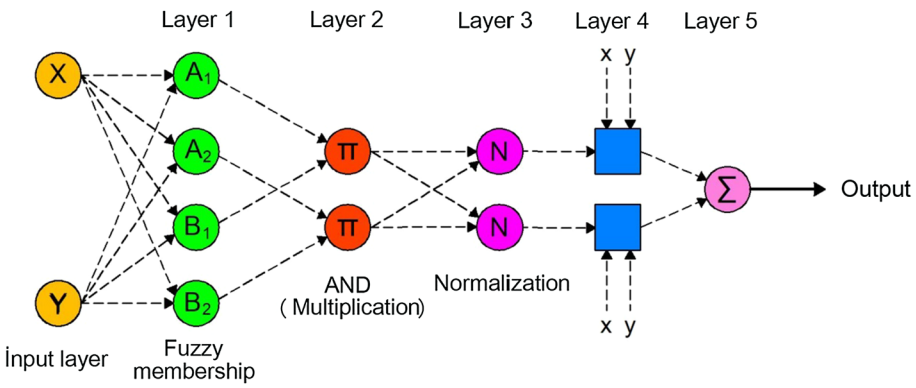

Fig. 1 ANFIS architecture structure (Tabari et al. 2012)

Layer 5: This layer is named the output nodes where the single node calculates the overall output considering the sum of all incoming signals as in Eq. (8) (Tabari et al. 2012):

$$O_i^5 = \sum_{i=1}^{2} \overline{w}_i z_i = \frac{w_1 z_1 + w_2 z_2}{w_1 + w_2}. \tag{8}$$

Consequently, the output z in Fig. 1 is obtained as:

$$Z = (\overline{w}_1 x)p_1 + (\overline{w}_1 y)q_1 + (\overline{w}_1)r_1 + (\overline{w}_2 x)p_2 + (\overline{w}_2 y)q_2 + (\overline{w}_2)r_2. \tag{9}$$

More information about ANFIS can be reached from the study of (Jang 1993). In this study, three ANFIS models, namely ANFIS-FCM, ANFIS-GP and ANFIS-SC were evaluated. ANFIS-GP model is obtained by combining ANFIS and grid partition. Grid partition approach separates the input section into rectangular subspaces utilizing a number of local fuzzy regions by axis-paralleled partition based on a predefined number of MFs and their types in each dimension (Sanikhani et al. 2012). ANFIS-SC approach is realized by combining ANFIS and subtractive clustering. It is an extension of that method in which each data point is assumed as a center for potential cluster center (Chiu 1994). On the other hand, ANFIS-FCM model is accomplished by combining ANFIS and Fuzzy C-Means (FCM) clustering method, which is a modification and improvement of K-means clustering (Bezdek 1981; Yager and Filev 1994).

2.2 Long short-term memory (LSTM) network

Deep learning, one of the fields of machine learning, enables machines to think and behave like humans. It uses a combination of several non-linear transformations to abstract high-level accuracy from the given dataset. Deep Neural Network (DNN), Convolutional Neural Network (CNN) and RNN are among the most commonly used deep learning models proposed to improve the results of conventional signal processing techniques in many applications (Park et al. 2019).

The LSTM network, a special kind of RNNs, was introduced by Hochreiter and Schmidhuber (1997). LSTM solves problems in traditional RNNs by adding cell states or memory cells with constant errors, so that errors can be reproduced without vanishing gradients. Input gate, forget gate and output gate are the three gates that take part in the LSTM network (Zahroh et al. 2019). Input gate learns to keep constant error flow in the core memory cell, from the irrelevant data. Output gate learns to keep other units from irrelevant core memory ingredients stored in core memory cells. Forget gate learns to check the endurance time of a dataset within the memory cell (Salman et al. 2018).

Hereafter, the dataset is divided into two categories as training and testing data. Training data are utilized for modeling and test data are used for model validation. The determination of parameters, such as maximum epoch, number of neurons in the hidden layer, number of hidden layers, batch size and learning rate may be random or algorithmic. Each gate function is calculated and executed until the maximum epoch or expected error target is reached (Zahroh et al. 2019). LSTM networks can be used for time-series prediction by learning long-term dependencies (Hochreiter and Schmidhuber 1997). The chains of recurrent LSTM units connected to each other comprise the LSTM network (Fig. 2).

In Fig. 2, the sequence data $x=(x_1, x_2, x_3, \ldots, x_t)$ refer to input variable to retrieve the hidden (output) state $h=(h_1, h_2, h_3, \ldots, h_t)$ and cell state $c=(c_1, c_2, c_3, \ldots, c_t)$. The first value of the sequence of x (x_1) is utilized in the first LSTM unit to obtain the first value of hidden state (h_1) and the first updated cell state (c_1). At time step t, c_{t-1} and h_{t-1} fed the LSTM unit to obtain h_t and c_t. The hidden state (h_t) at time t with cell state (c_t) can be calculated by:

$$h_t = o_t \odot \tanh\left(c_t\right), \tag{10}$$

where $\odot$ represents the Hadamard product (element-wise multiplication of vectors). o_t refers to the output gate and controls the level of the cell state added to hidden state. The cell state manages the LSTM network by adding or removing information to it using the gates. The cell state (c_t) the time step t consists of information from the previous units and can be computed as follows:

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t. \tag{11}$$

In Eq. (11), input gate (i_t) controls the level of cell state update. Forget gate (f_t) manages the level of cell state reset. Cell candidate (g_t) adds information to the cell state. The variables i_t, f_t, g_t, and o_t can be computed using the following equations, respectively.

$$i_t = \sigma\left(W_i x_t + R_i h_{t-1} + b_i\right), \tag{12}$$

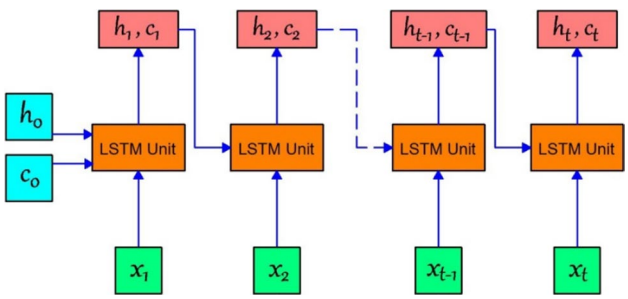

Fig. 2 The architecture of LSTM network

$$f_t = \sigma\left(W_f x_t + R_f h_{t-1} + b_f\right), \tag{13}$$

$$g_t = \tanh\left(W_g x_t + R_g h_{t-1} + b_g\right), \tag{14}$$

$$o_t = \sigma\left(W_o x_t + R_o h_{t-1} + b_o\right), \tag{15}$$

where σ refers to sigmoid function obtained by $\sigma(x) = (1 + e^{-x})^{-1}$. W is the input weights. R represents the recurrent weights and b is the bias term. All of these parameters can be given by Eq. (16).

$$W = \begin{bmatrix} W_i \\ W_f \\ W_g \\ W_o \end{bmatrix}, \quad R = \begin{bmatrix} R_i \\ R_f \\ R_g \\ R_o \end{bmatrix}, \quad b = \begin{bmatrix} b_i \\ b_f \\ b_g \\ b_o \end{bmatrix}. \tag{16}$$

Each LSTM unit in Fig. 2 can be represented by the architecture in Fig. 3 (Mathworks 2019).

2.3 Error analysis

In this study, three statistical error criteria were considered to evaluate the efficiency of the models in estimating an observed output variable. The three statistical metrics can be specified by:

Root Mean square error (RMSE)

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[p(i) - o(i)\right]^2}, \tag{17}$$

Correlation Coefficient (R),

$$R = \left(\sum_{i=1}^{N}\left[p(i) - \overline{p}\right]\left[o(i) - \overline{o}\right]\right) \Big/ \left(\sqrt{\sum_{i=1}^{N}\left[p(i) - \overline{p}\right]^2}\sqrt{\sum_{i=1}^{N}\left[o(i) - \overline{o}\right]^2}\right), \tag{18}$$

and Mean Absolute Error (MAE)

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|p(i) - o(i)|, \tag{19}$$

where $o(i)$ and $p(i)$ refer to the observed value and predicted value at a certain time i, respectively. On the other hand, $\overline{o}$ and $\overline{p}$ stand for the average observed values and the average predicted values, respectively. N is the total number of the data.

3 Study area, data and results

3.1 Study area and data analysis

The city of Tarsus is one of the districts of Mersin province in Turkey, and its immediate surroundings are dominated by Mediterranean climatic conditions. The average annual temperature in the city is 18 °C. The lowest monthly average temperature is observed as 9 °C in January, while the highest is 27 °C in July. The most fertile and the widest agricultural land in Mersin province is in the region of Tarsus Plain. The climate is suitable for agriculture in those lands provide all kinds of agricultural production. Tarsus district has 104,902 hectares of farmland. 68 different kinds of agricultural products are grown commercially in this region. In addition to its agricultural fertility, Tarsus has the highest solar radiation energy source with approximately 1700 kWh/m^2 in a year. Therefore, the installation of solar power plants in this region has been increased in recent years. In this study, the measured AT data were obtained at hourly and daily intervals from a solar power plant with a capacity of 1 MW. The observation range of the temperature measurement device is between − 50 and 70 °C, and its accuracy is ± 0.10 °C. The geographic coordinates of the solar power plant are 37° 07′ 19″ N and 34° 45′ 37″ E. Figure 4a shows the general overview of the study area and Fig. 4b illustrates the solar power plant station where the measurements are taken.

In this paper, the ANFIS model was categorized into two parts including construction and training part. The type and number of MFs were specified in the construction part. Considering the ANFIS model, it is required to divide the input/output data into rule patches. To train an ANFIS model, training data pairs were first determined in the training part.

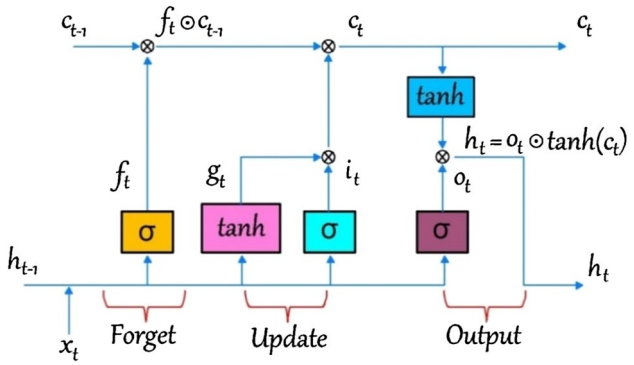

Fig. 3 The architecture of each LSTM unit

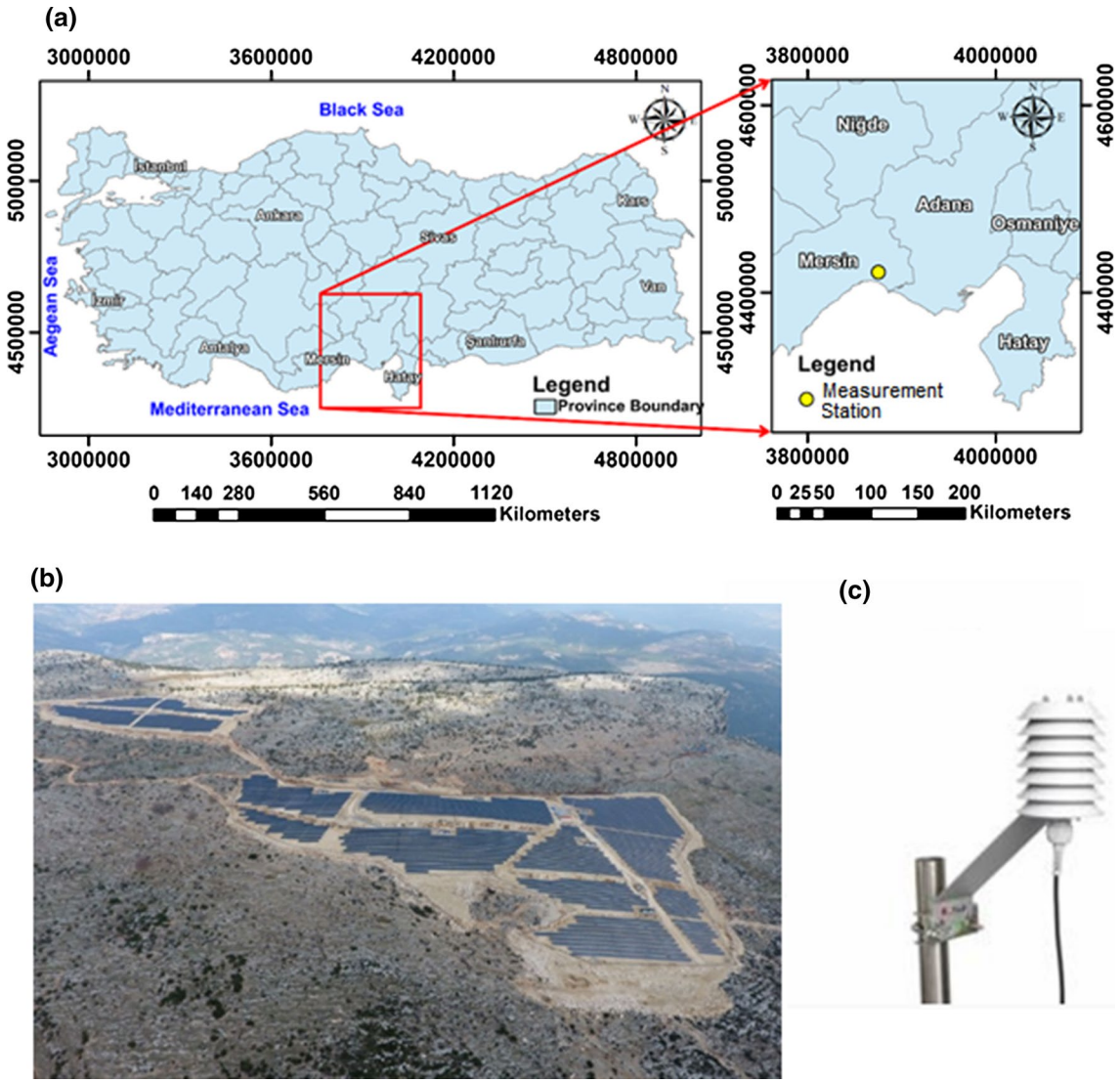

Fig. 4 The study area **a** general overview of the measurement station in Turkey, **b** view of the solar power plant, **c** image of temperature measurement device installed in the plant

These data pairs represent the inputs and related output of the ANFIS model. Three different methods (Yarpiz 2020), namely, ANFIS-FCM, ANFIS-SC and ANFIS-GP were applied to perform one-hour ahead short-term AT prediction. We used Sugeno's fuzzy approach for the retrieval of output variable from the input variables given to the FIS structure. Considering different ANFIS architectures, the most suitable model structures, number of MFs, number of iterations (epoch number) and transfer functions were identified using trial and error method. In the construction part, each ANFIS model was tested and the statistical metrics were utilized to compare their performances. Gaussmf and linear membership functions were used as input and output MFs, respectively. The number of MFs ranged from 2 to 6, and the number of iterations varied between 50 and 300 in the processing phase.

In the ANFIS and LSTM model simulations, AT data were divided into training and testing datasets to train and validate the models, respectively. The RMSE, MAE and R were used as evaluation criteria. The one-day intervals AT data series, from January 1, 2018 to December 31, 2019, were utilized for one-day ahead short-term AT prediction. The total 730 obtained samples were divided into two parts (80% as the training set and 20% as the test set). On the

Table 1 The hourly prediction performances of each model

Model	Properties	MAE (°C)	RMSE (°C)	*R*
ANFIS-FCM	Number of MFs: 2	0.5162	0.7237	0.9790
ANFIS-FCM	Number of MFs: 3	0.5186	0.7233	0.9790
ANFIS-FCM	Number of MFs: 4	0.5184	0.7217	0.9791
ANFIS-FCM	**Number of MFs: 5**	**0.5164**	**0.7206**	**0.9792**
ANFIS-FCM	Number of MFs: 6	0.5160	0.7209	0.9792
ANFIS-SC	Influence radius: 0.2	0.5222	0.7267	0.9778
ANFIS-SC	**Influence radius: 0.3**	**0.5176**	**0.7216**	**0.9791**
ANFIS-SC	Influence radius: 0.5	0.5187	0.7236	0.9789
ANFIS-SC	Influence radius: 0.7	0.5194	0.7240	0.9789
ANFIS-GP	**Number of MFs: 2**	**0.5668**	**0.8301**	**0.9723**
ANFIS-GP	Number of MFs: 3	0.7005	1.1985	0.9408
LSTM	Number of hidden layer: 20	0.5322	0.7255	0.9782
LSTM	Number of hidden layer: 60	0.5451	0.7495	0.9786
LSTM	Number of hidden layer: 100	0.5105	0.7043	0.9807
LSTM	Number of hidden layer: 150	0.4950	0.6825	0.9804
LSTM	**Number of hidden layer: 200**	**0.4597**	**0.6567**	**0.9823**
LSTM	**Number of hidden layer: 250**	**0.4597**	**0.6436**	**0.9827**

Best results are shown in bold

other hand, the hourly AT data recorded from October 4, 2014 to June 24, 2015 were used to perform one-hour ahead short-term AT prediction. The total of 6309-h data was split into two parts, first as 80% training set and the other 20% testing set.

3.2 Prediction results for the hourly AT data

Table 1 represents different accuracy measures for both ANFIS models and LSTM network. Concerning the ANFIS-FCM method, the best performance was obtained from the MFs equal to 5 with the values of 0.5164 °C MAE, 0.7206 °C RMSE and 0.9792 *R*. Influence radius of 0.3 in the ANFIS-SC method gave the best output values of 0.5176 °C MAE, 0.7216 °C RMSE and 0.9791 *R*. Considering the ANFIS-GP method, the best results were observed when the MFs are equal to 2 with the values of 0.5668 °C MAE, 0.8301 °C RMSE and 0.9723 *R*. The results showed that all ANFIS models presented satisfying performances in hourly AT prediction. However, ANFIS-FCM and ANFIS-SC models gave almost identical results which revealed that they were better than ANFIS-GP.

Apart from ANFIS models, LSTM method was applied for different numbers of hidden layers as in Table 1. The hidden layers of 200 and 250 presented almost similar results in terms of accuracy measures; however, the hidden layer of 250 provided slightly better RMSE and *R*. For the hidden layer number equal to 200, MAE, RMSE and *R* were obtained as 0.4597 °C, 0.6567 °C and 0.9823, respectively. Concerning the hidden layer number equal to 250, MAE, RMSE and *R* were calculated as 0.4597 °C, 0.6436 °C and 0.9827, respectively. The RMSE values from the all methods (ANFIS models and LSTM) indicated that the best results were acquired from LSTM method with the values 0.6567 (hidden layer: 200) and 0.6436 (hidden layer: 250), respectively. As the R measure was examined for all methods, they gave similar results about 0.98. This phenomenon shows the quality of fitting models to data for prediction. Consequently, the results of the statistical metrics in Table 1 showed that the LSTM method presented more favorable accuracies compared to the ANFIS methods.

Fig. 5a–d shows training and testing time series with observed and predicted hourly AT data for ANFIS-FCM, ANFIS-SC, ANFIS-GP and LSTM methods, respectively. *X* axis shows the number of hourly samples while *Y* axis corresponds to AT in °C. The hourly temperature changes can be seen in AT time series. The graphic in Fig. 5 show that the prediction of the AT time series coincides with the observed values in testing part for all the methods. This situation was also proved by the statistical metrics in Table 1. ANFIS models and LSTM network provided satisfactory results with this kind of sinusoidal data. To analyze the prediction results in more detail, Fig. 6 is presented as a close look to testing values. Figure 6 proves that the LSTM network made more reliable prediction for maximum and minimum peak values of the data compared to all ANFIS models.

In addition to Figs. 6, 7a–d shows regression plots of observed and predicted values of AT data from

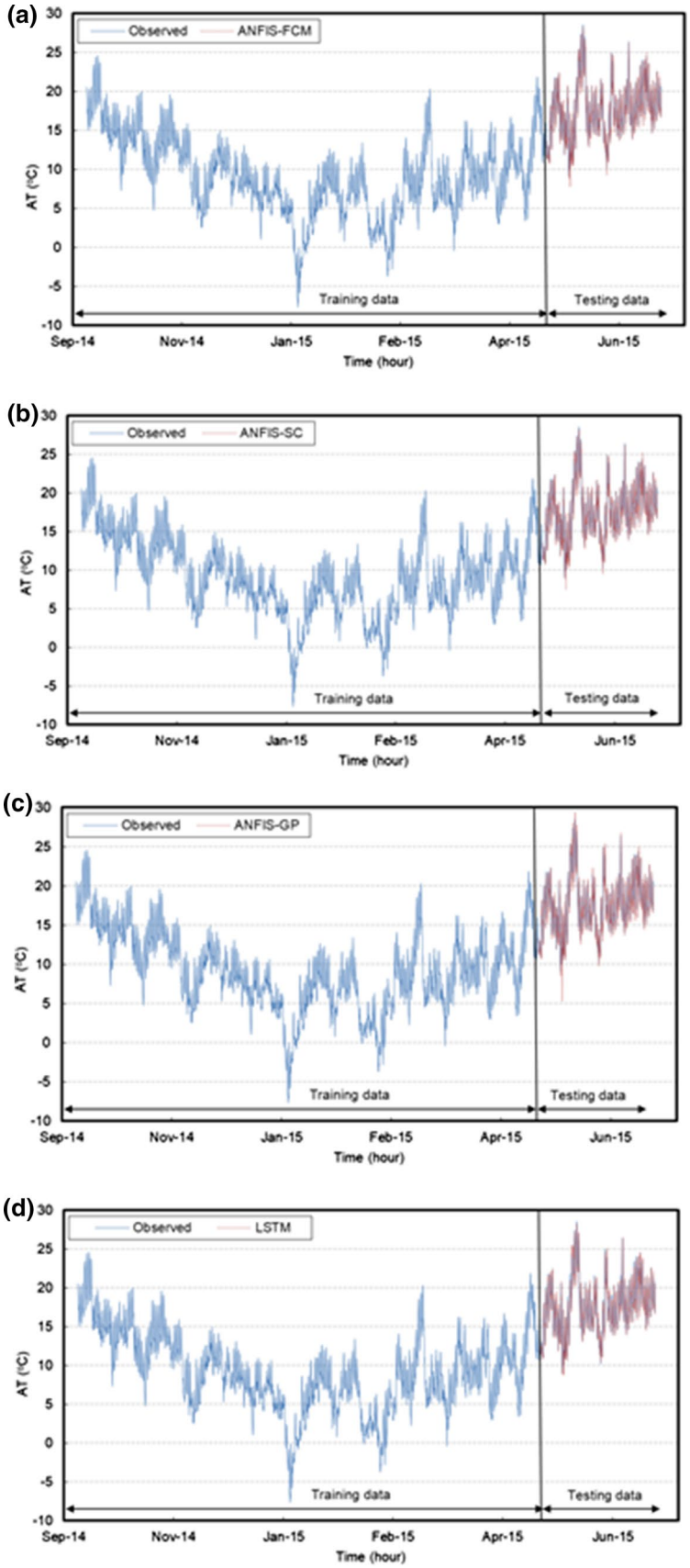

Fig. 5 The hourly time-series AT data of Tarsus Station with observed (blue) and predicted values (red) for **a** ANFIS-FCM, **b** ANFIS-SC **c** ANFIS-GP and **d** LSTM methods

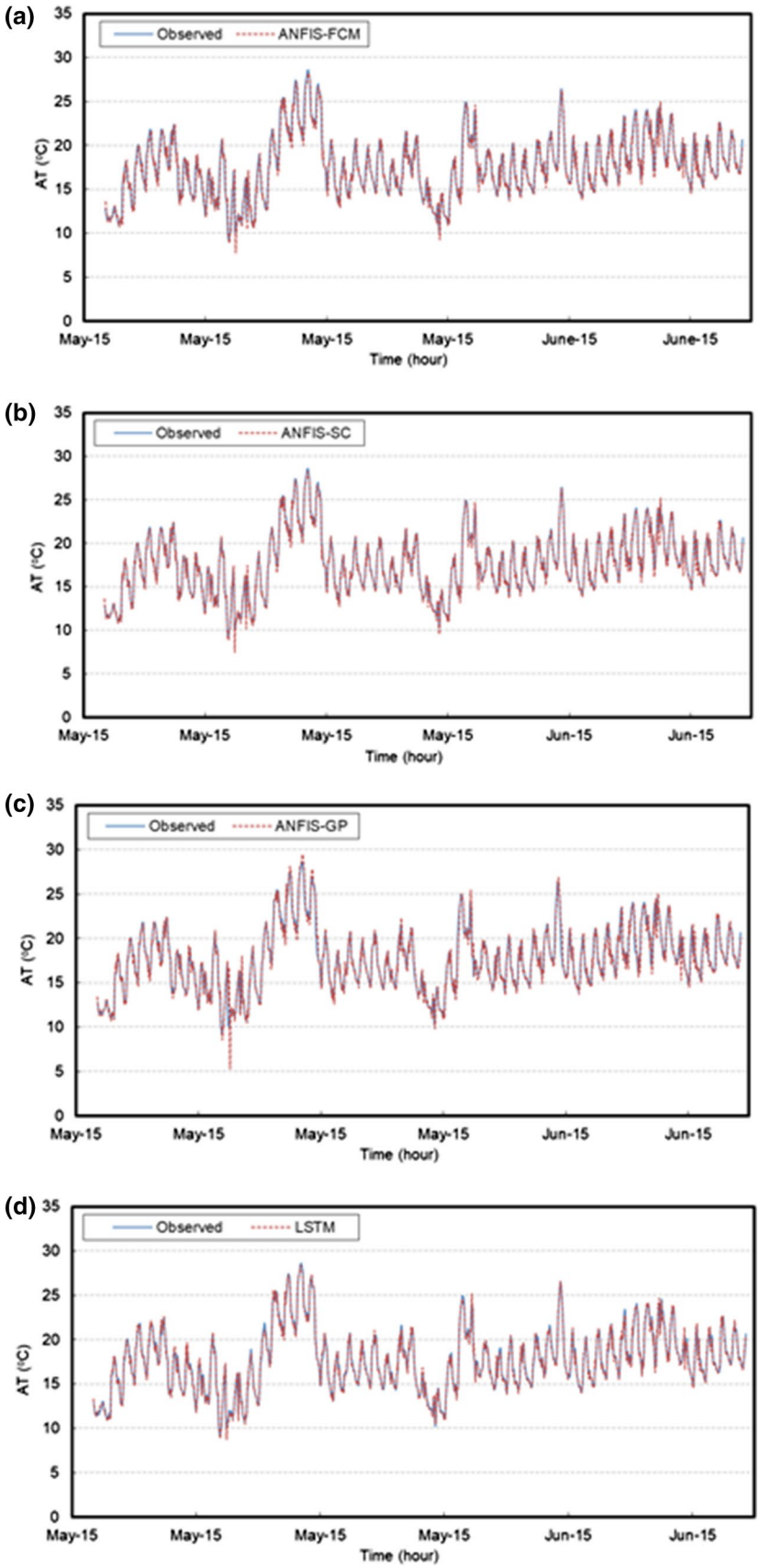

Fig. 6 Observed (blue) and predicted values (red) of testing hourly AT data for Tarsus station from **a** ANFIS-FCM, **b** ANFIS-SC **c** ANFIS-GP and **d** LSTM methods

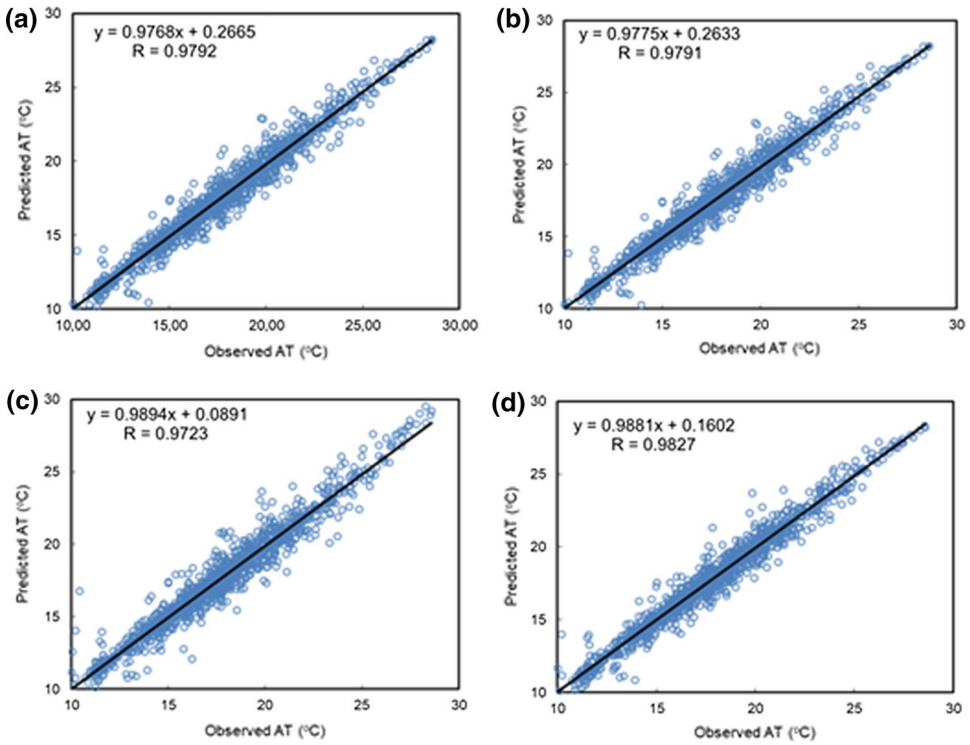

Fig. 7 Regression plots of the observed and predicted values of AT data for Tarsus station from **a** ANFIS-FCM, **b** ANFIS-SC **c** ANFIS-GP and **d** LSTM methods

Table 2 The daily prediction performances of each model

Model	Properties	MAE (°C)	RMSE (°C)	*R*
ANFIS-FCM	**Number of MFs: 2**	**1.0263**	**1.3660**	**0.9691**
ANFIS-FCM	Number of MFs: 3	1.0608	1.4325	0.9659
ANFIS-FCM	Number of MFs: 4	1.0771	1.4278	0.9659
ANFIS-FCM	Number of MFs: 5	1.0846	1.4460	0.9651
ANFIS-FCM	Number of MFs: 6	1.0792	1.4465	0.9651
ANFIS-SC	Influence radius: 0.2	1.0607	1.4105	0.9669
ANFIS-SC	Influence radius: 0.3	1.0886	1.4593	0.9644
ANFIS-SC	**Influence radius: 0.5**	**1.0401**	**1.4052**	**0.9672**
ANFIS-SC	Influence radius: 0.7	1.0452	1.4105	0.9670
ANFIS-GP	**Number of MFs: 2**	**1.2456**	**1.9054**	**0.9406**
LSTM	**Number of hidden layer: 5**	**0.9939**	**1.3599**	**0.9701**
LSTM	Number of hidden layer: 20	0.9925	1.3719	0.9705
LSTM	Number of hidden layer: 60	1.0463	1.3962	0.9681
LSTM	Number of hidden layer: 100	1.0453	1.4023	0.9680
LSTM	Number of hidden layer: 150	1.1698	1.5906	0.9578
LSTM	Number of hidden layer: 200	1.4789	2.2358	0.9219
LSTM	Number of hidden layer: 250	1.2992	1.8040	0.9457

Best results are shown in bold

ANFIS-FCM, ANFIS-SC, ANFIS-GP and LSTM methods, respectively, for Tarsus station. *X* axis shows observed (actual) AT data, while *Y* axis corresponds to predicted AT in °C. The graphs in Fig. 7 show the distribution of the observed and predicted values, making it possible to understand how accurately the model results fit to actual data. For ANFIS-FCM, ANFIS-SC, ANFIS-GP methods, *R* values are 0.9792, 0.9791 and 0.9723, respectively. As it is clear from *R* value of 0.9827, the model fit is better for LSTM method compared to all ANFIS models.

3.3 Prediction results for the daily AT data

Table 2 shows the statistical performance metrics of the daily AT forecasting for all models with different properties (parameters). The number of MFs were varied between 2 and 6 for the ANFIS-FCM method. The quality metrics varied from 1.026 to 1.079 °C (MAE), 1.366 to 1.447 °C (RMSE) and 0.9651 to 0.9691 (R), respectively. The best obtained RMSE is 1.366 °C for MF = 2 in ANFIS-FCM. Considering the ANFIS-SC method, the influence radius varied between 0.2 and 0.7. The quality metrics were ranged from 1.040 °C to 1.089 °C (MAE), 1.405 to 1.459 °C (RMSE) and 0.9644 to 0.9672 (R), respectively. The best obtained RMSE was 1.405 °C in ANFIS-SC (influence radius equal to 0.5). On the other hand, the statistical measures were 1.246 °C (MAE), 1.905 °C (RMSE) and 0.9406 (R), respectively, for the best result of the ANFIS-GP method (MF = 2). As in the hourly AT prediction, ANFIS-FCM presented the better results than ANFIS-SC and ANFIS-GP.

The selected hidden layer number in LSTM network ranged from 5 to 250. The statistical measures of LSTM models were varied from 0.994 to 1.479 °C (MAE), 1.360 to 2.236 °C (RMSE) and 0.9219 to 0.9701 (R), respectively. The best performance was achieved using the hidden layer number as 5 with the RMSE value of 1.360 °C. It can be said that if the hidden layer size more increased, the solution does not get higher accuracy, and the RMSE values are getting higher. Thus, it is crucial to select optimal parameters to get high accuracy results from both ANFIS and LSTM network. Besides, another important result in Table 2 is that some of the LSTM models provided higher accuracy than any ANFIS model.

Figure 8a–d depicts training and testing data with observed and forecasted daily AT data for ANFIS-FCM, ANFIS-SC, ANFIS-GP and LSTM methods, respectively. 585 sample data were utilized as the training data, whereas 145 sample data as the testing data. The observed and forecasted values are shown in blue and red, respectively. From Fig. 8a–d, it can be seen that the forecasted AT data almost overlap with the observed values in testing part for all the methods. However, the LSTM values give better forecast results that can be seen as a lower value in RMSE. To visualize the quality of the forecasting in more detail, Fig. 9a–d is just presented for the testing data.

To show one of the quality metrics of fitting model to data as correlation coefficient, the regression plot of observed and forecasted daily AT data is represented in Fig. 10. Figure 10a–d shows regression plot of ANFIS-FCM, ANFIS-SC, ANFIS-GP and LSTM methods, respectively. The correlation coefficient (R) for four methods can be given as 0.9701 (LSTM), 0.9691 (ANFIS-FCM), 0.9672 (ANFIS-SC) and 0.9406 (ANFIS-GP), respectively. As it can be seen from the figures, the best model fit to data is carried out by the LSTM model followed by the ANFIS-FCM method.

3.4 Comparison of the hourly and daily AT predictions

To sum up, the statistical accuracy results of the hourly and daily AT predictions are emphasized in Table 3. As pointed out in the previous sections, LSTM provided the best accuracies for both daily and hourly AT prediction compared to ANFIS models. On the other hand, ANFIS-FCM and ANFIS-SC was provided close accuracy results which were better than ANFIS-GP. Considering the comparison of hourly and daily AT prediction with LSTM method, it is clear from Table 3 that LSTM modeled and predicted the hourly data better than the daily data. Moreover, ANFIS models behaved in the same manner like LSTM with regard to hourly and daily data. This is most probably due to the fact that hourly data does not have instantaneous variations compared to daily data. Therefore, the prediction accuracies of all models decreased in daily AT prediction; nevertheless, the best results for both LSTM and ANFIS models are still satisfactory.

4 Conclusion

In this study, one-hour and one-day ahead prediction of AT time series were performed using machine learning methods of ANFIS-FCM, ANFIS-SC, ANFIS-GP and LSTM network. The total of 6309 hourly data and 730 daily data were split into two parts as 80% training set and 20% testing set for all methods. The results revealed that 80% of the training

Table 3 Statistical accuracy results of the hourly and daily AT prediction

Forecasting interval	Forecasting method	Error criteria		
		MAE (°C)	RMSE (°C)	R
Hourly	**LSTM**	**0.4597**	**0.6436**	**0.9827**
	ANFIS-FCM	0.5164	0.7206	0.9792
	ANFIS-SC	0.5176	0.7216	0.9791
	ANFIS-GP	0.5668	0.8301	0.9723
Daily	**LSTM**	**0.9939**	**1.3599**	**0.9701**
	ANFIS-FCM	1.0263	1.3660	0.9691
	ANFIS-SC	1.0401	1.4052	0.9672
	ANFIS-GP	1.2456	1.9054	0.9406

The best results are shown in bold

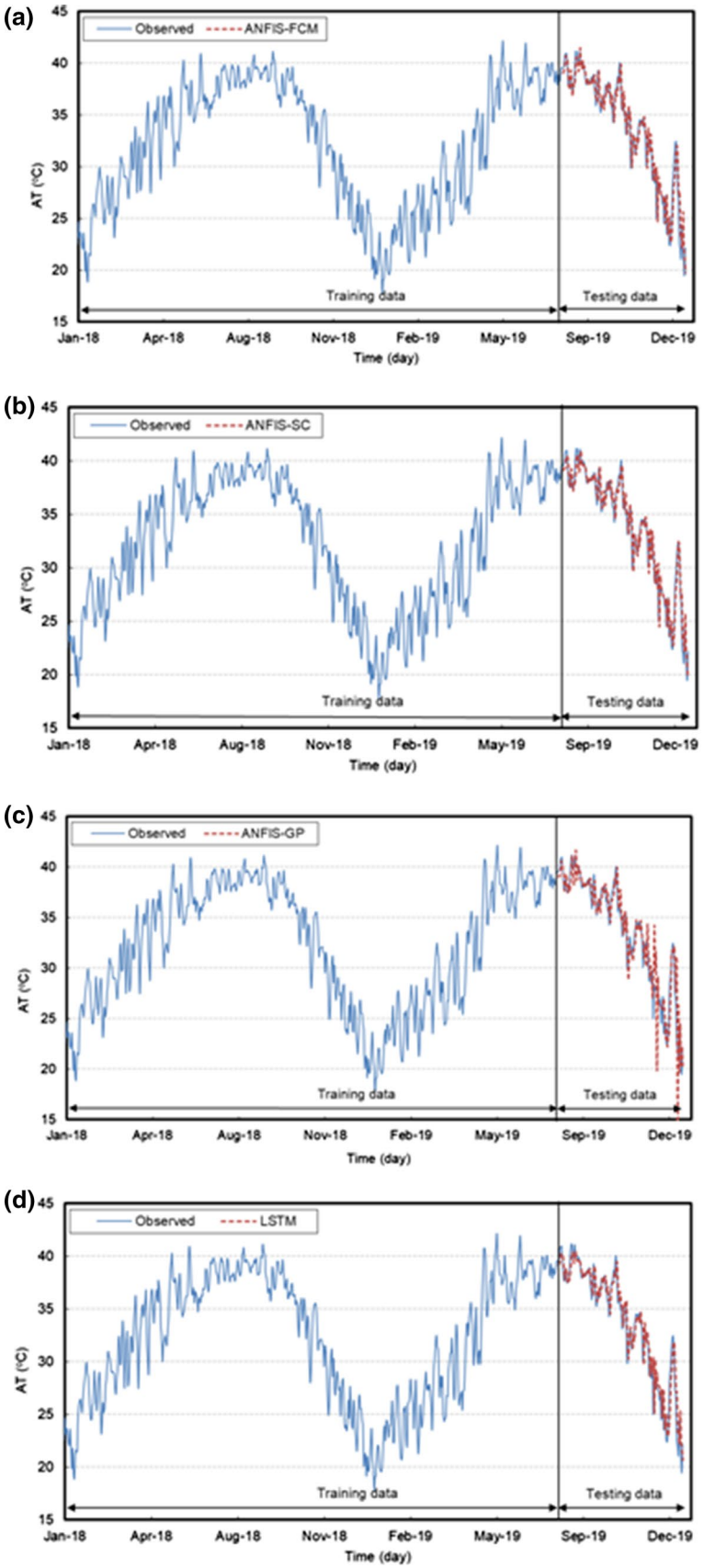

Fig. 8 The daily time-series AT data of Tarsus Station with observed (blue) and predicted values (red) for **a** ANFIS-FCM, **b** ANFIS-SC **c** ANFIS-GP and **d** LSTM methods

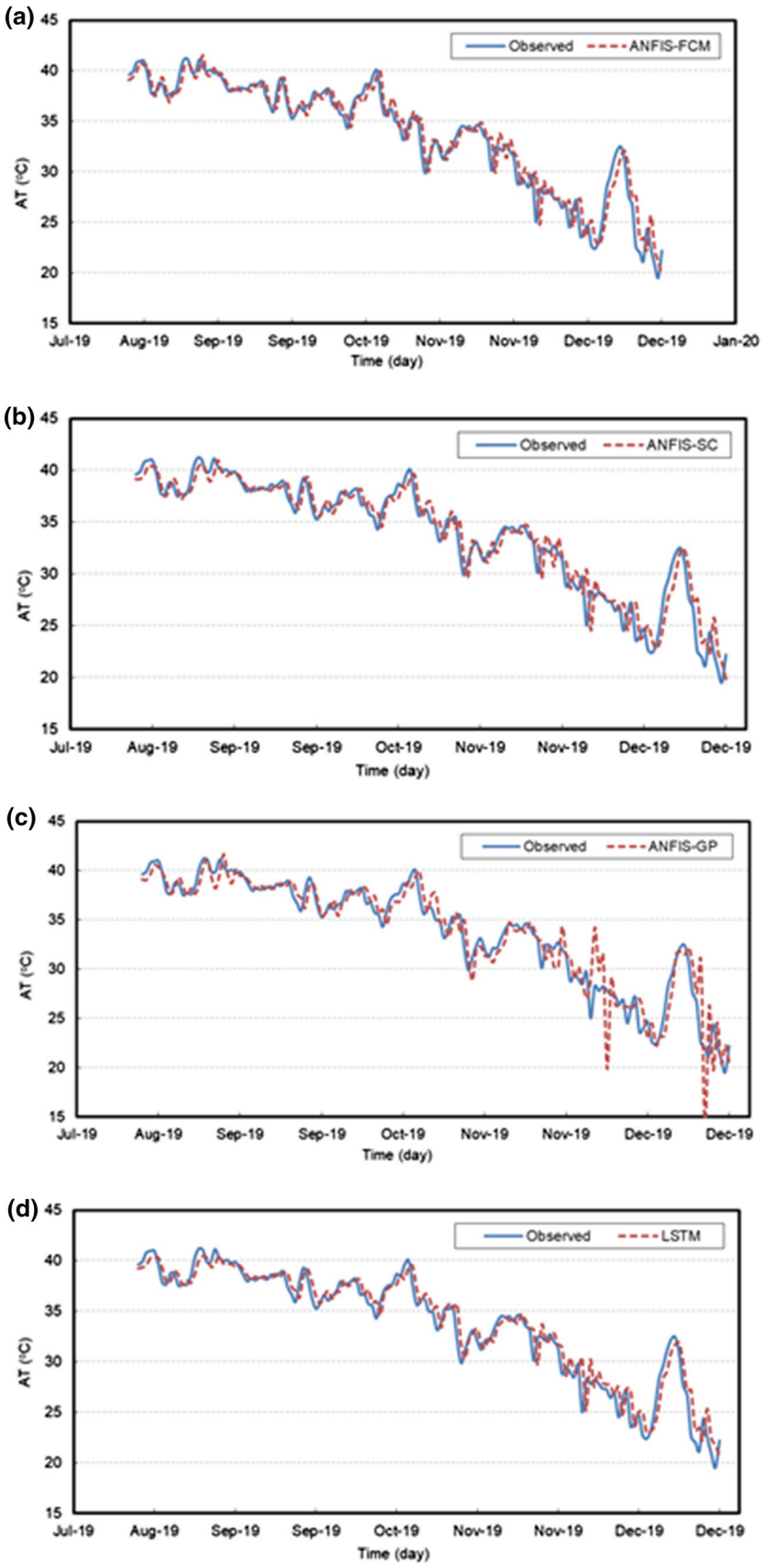

Fig. 9 Observed (blue) and predicted values (red) of testing daily AT data for Tarsus station from **a** ANFIS-FCM, **b** ANFIS-SC **c** ANFIS-GP and **d** LSTM methods

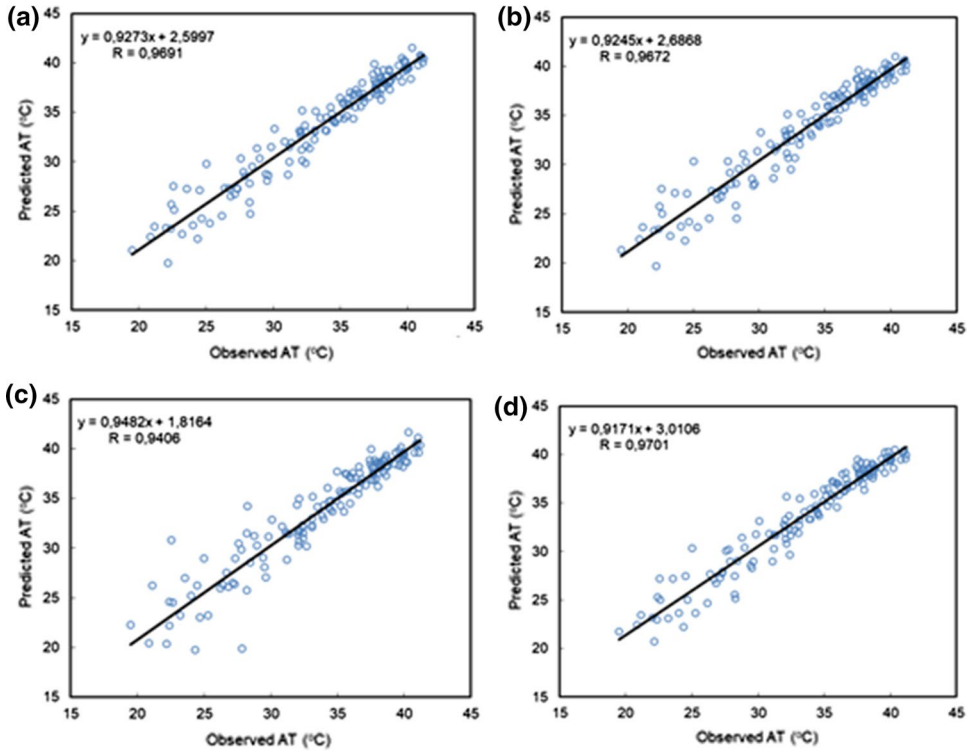

Fig. 10 Regression plots of the observed and predicted values of AT data for Tarsus station from **a** ANFIS-FCM, **b** ANFIS-SC **c** ANFIS-GP and **d** LSTM methods

data was sufficient for the reliable prediction of hourly and daily AT using the corresponding methods. The optimal input parameters such as number of MFs, number of iterations (epoch number) and transfer functions were identified using trial and error for all methods. All ANFIS models used in this study presented satisfying performances in hourly and daily AT predictions. However, the inter-comparison of the ANFIS models indicated that ANFIS-FCM and ANFIS-SC models provided almost identical results, slightly better than ANFIS-GP. On the other hand, LSTM network provided the best accuracy results amongst all methods, showing its advantage and capability to learn long-term dependency. The RMSE values of LSTM (0.6436–1.3599), ANFIS-GP (0.8301–1.9054), ANFIS-SC (0.7216–1.4052) and ANFIS-FCM (0.7206–1.3660) for hourly and daily predictions, respectively, proved the highest performance of LSTM compared to the ANFIS models. Considering the comparison of hourly and daily AT prediction results, all methods modeled and predicted the hourly data better than the daily data, which is most probably due to the fact that hourly data do not have sudden changes compared to daily data. However, the best results for both LSTM and ANFIS models are still satisfactory. Overall, this study has demonstrated that different ANFIS models and LSTM network are powerful tools in predicting station-based hourly and daily AT data.

Acknowledgements We would like to thank Pilye Energy Cons. Ind. and Trading Co., who most kindly and generously allowed us to use the solar power plant temperature data in our study.

References

Abyaneh HZ, Nia AM, Varkeshi MB, Marofi S, Kisi O (2011) Performance evaluation of ANN and ANFIS models for estimating garlic crop evapotranspiration. J Irrig Drain Eng 137:280–286. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000298

Arslan N, Sekertekin A (2019) Application of long short-term memory neural network model for the reconstruction of MODIS land surface temperature images. J Atmos Solar Terr Phys 194:105100. https://doi.org/10.1016/j.jastp.2019.105100

Azad A, Kashi H, Farzin S, Singh VP, Kisi O, Karami H, Sanikhani H (2019) Novel approaches for air temperature prediction: a comparison of four hybrid evolutionary fuzzy models. Meteorol Appl. https://doi.org/10.1002/met.1817

Bandara TMDK, Yapa RD, Kodituwakku SR (2011) Simulation of regression analysis by an automated system utilizing artificial neural networks . Int J Latest Trends Comput 2:378–391

Benmouiza K, Cheknane A (2019) Clustered ANFIS network using fuzzy c-means, subtractive clustering, and grid partitioning for hourly solar radiation forecasting. Theor Appl Climatol 137:31–43. https://doi.org/10.1007/s00704-018-2576-4

Bezdek JC (1981) Pattern recognition with fuzzy objective function algorithms. Springer, 272 p. https://doi.org/10.1007/978-1-4757-0450-1

Bilgili M (2010) Prediction of soil temperature using regression and artificial neural network models. Meteorol Atmos Phys 110:59–70. https://doi.org/10.1007/s00703-010-0104-x

Bilgili M, Sahin B (2010) Prediction of long-term monthly temperature and rainfall in Turkey. Energy sources, part A recover. Util Environ Eff 32:60–71. https://doi.org/10.1080/15567030802467522

Capone M (2020) Predictive Analytics [WWW Document]. https://www.qlik.com/us/predictive-analytics. Accessed 27 Dec 20

Celebi K, Uludamar E, Tosun E, Yildizhan S, Aydin K, Ozcanli M (2017) Experimental and artificial neural network approach of noise and vibration characteristic of an unmodified diesel engine

fuelled with conventional diesel, and biodiesel blends with natural gas addition. Fuel 197:159–173. https://doi.org/10.1016/j.fuel.2017.01.113

Chaudhuri S, Middey A (2011) Adaptive neuro-fuzzy inference system to forecast peak gust speed during thunderstorms. Meteorol Atmos Phys 114:139. https://doi.org/10.1007/s00703-011-0158-4

Chen J, Zeng GQ, Zhou W, Du W, Di LuK (2018) Wind speed predicting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization. Energy Convers Manag 165:681–695. https://doi.org/10.1016/j.enconman.2018.03.098

Chevalier RF, Hoogenboo G, McClendon RW, Paz JA (2011) Support vector regression with reduced training sets for air temperature prediction: a comparison with artificial neural networks. Neural Comput Appl 20:151–159. https://doi.org/10.1007/s00521-010-0363-y

Chiu SL (1994) Fuzzy model identification based on cluster estimation. J Intell Fuzzy Syst 2(3):267–278. https://doi.org/10.3233/IFS-1994-2306

Christensen JH, Hewitson B, Busuioc A, Chen A, Gao X, Held I, Jones R, Kolli RK, Kwon WT, Laprise R, Rueda VM, Mearns L, Menéndez CG, Räisänen J, Rinke A, Sarr A, Whetton P (2007) Regional climate projections. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) Climate change. The physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge

Cobaner M, Citakoglu H, Kisi O, Haktanir T (2014) Estimation of mean monthly air temperatures in Turkey. Comput Electron Agric 109:71–79. https://doi.org/10.1016/j.compag.2014.09.007

Han S, Qiao YH, Yan J, Liu YQ, Li L, Wang Z (2019) Mid-to-long term wind and photovoltaic power generation prediction based on copula function and long short term memory network. Appl Energy 239:181–191. https://doi.org/10.1016/j.apenergy.2019.01.193

Hewage P, Trovati M, Pereira E, Behera A (2020) Deep learning-based effective fine-grained weather forecasting model. Pattern Anal Appl. https://doi.org/10.1007/s10044-020-00898-1

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9:1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hoogenboom G (2000) Contribution of agrometeorology to the simulation of crop production and its applications. Agric For Meteorol. https://doi.org/10.1016/S0168-1923(00)00108-8

Jang JR (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23:665–685

Karakuş O, Kuruoğlu EE, Altinkaya MA (2017) One-day ahead wind speed/power prediction based on polynomial autoregressive model. IET Renew Power Gener 11:1430–1439. https://doi.org/10.1049/iet-rpg.2016.0972

Kumar S, Roshni T, Kahya E, Ghorbani MA (2020) Climate change projections of rainfall and its impact on the cropland suitability for rice and wheat crops in the Sone river command, Bihar. Theor Appl Climatol 142(1):433–451. https://doi.org/10.1007/s00704-020-03319-9

Li X, Li Z, Huang W, Zhou P (2020) Performance of statistical and machine learning ensembles for daily temperature downscaling. Theor Appl Climatol 140:571–588. https://doi.org/10.1007/s00704-020-03098-3

Liu H, Mi X, Li Y (2018) Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolutional neural network and convolutional long short term memory network. Energy Convers Manag 166:120–131. https://doi.org/10.1016/j.enconman.2018.04.021

Liu H, He B, Qin P, Zhang X, Guo S, Mu X (2020) Sea level anomaly intelligent inversion model based on LSTM-RBF network. Meteorol Atmos Phys. https://doi.org/10.1007/s00703-020-00745-2

Ma X, Tao Z, Wang Y, Yu H, Wang Y (2015) Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. Transp Res Part C Emerg Technol 54:187–197. https://doi.org/10.1016/j.trc.2015.03.014

Mathworks (2019) Long Short-Term Memory Networks [WWW Document]. https://www.mathworks.com/help/deeplearning/ug/long-short-term-memory-networks.html

Meshram SG, Kahya E, Meshram C, Ghorbani MA, Ambade B, Mirabbasi R (2020) Long-term temperature trend analysis associated with agriculture crops. Theoret Appl Climatol 140:1139–1159. https://doi.org/10.1007/s00704-020-03137-z

Misra S, Sarkar S, Mitra P (2018) Statistical downscaling of precipitation using long short-term memory recurrent neural networks. Theor Appl Climatol 134:1179–1196. https://doi.org/10.1007/s00704-017-2307-2

Nag PK, Nag A, Sekhar P, Pandit S (2009) Vulnerability to heat stress: Scenario in Western India. (WHO report APW No. SO 08 AMS 6157206). World Health Organization Regional Office for South-East Asia, New Delhi

Park I, Kim HS, Lee J, Kim JH, Song CH, Kim HK (2019) Temperature prediction using the missing data refinement model based on a long short-term memory neural network. Atmosphere (Basel) 10:1–16. https://doi.org/10.3390/atmos10110718

Peng L, Liu S, Liu R, Wang L (2018) Effective long short-term memory with differential evolution algorithm for electricity price prediction. Energy 162:1301–1314. https://doi.org/10.1016/j.energy.2018.05.052

Qin Y, Li K, Liang Z, Lee B, Zhang F, Gu Y, Zhang L, Wu F, Rodriguez D (2019) Hybrid predicting model based on long short term memory network and deep learning neural network for wind signal. Appl Energy 236:262–272. https://doi.org/10.1016/j.apenergy.2018.11.063

Radhika Y, Shashi M (2009) Atmospheric temperature prediction using support vector machines. Int J Comput Theory Eng 1:55–58. https://doi.org/10.7763/ijcte.2009.v1.9

Ramesh K, Anitha R (2014) MARSpline model for lead seven-day maximum and minimum air temperature prediction in Chennai, India. J Earth Syst Sci 123:665–672. https://doi.org/10.1007/s12040-014-0434-z

Salman AG, Heryadi Y, Abdurahman E, Suparta W (2018) Single layer & multi-layer long short-term memory (LSTM) model with intermediate variables for weather predicting. Procedia Comput Sci 135:89–98. https://doi.org/10.1016/j.procs.2018.08.153

Sanikhani H, Kisi O, Nikpour MR, Dinpashoh Y (2012) Estimation of daily pan evaporation using two different adaptive neuro-fuzzy computing techniques. Water Resour Manage 26(15):4347–4365. https://doi.org/10.1007/s11269-012-0148-4

Scher S (2020) Artificial intelligence in weather and climate prediction. Stockholm University, Stockholm, Sweden, p. 30

Sekula P, Bokwa A, Bochenek B, Zimnoch M (2019) Prediction of air temperature in the Polish Western Carpathian Mountains with the ALADIN-HIRLAM numerical weather prediction system. Atmosphere (Basel). https://doi.org/10.3390/atmos10040186

Tabari H, Kisi O, Ezani A, Hosseinzadeh Talaee P (2012) SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment. J Hydrol 444–445:78–89. https://doi.org/10.1016/j.jhydrol.2012.04.007

Venkadesh S, Hoogenboom G, Potter W, McClendon R (2013) A genetic algorithm to refine input data selection for air temperature prediction using artificial neural networks. Appl Soft Comput J 13:2253–2260. https://doi.org/10.1016/j.asoc.2013.02.003

Yager RR, Filev DP (1994) Approximate clustering via the mountain method. IEEE Trans Syst Man Cybern 24(8):1279–1284. https://doi.org/10.1109/21.299710

Yarpiz (2020) Time-Series Prediction using ANFIS. (https://www.mathworks.com/matlabcentral/fileexchange/52969-time-series-prediction-using-anfis), MATLAB Central File Exchange. Retrieved 5 April 2020

Zahroh S, Hidayat Y, Pontoh RS, Santoso A, Sukono Bon AT (2019) Modeling and predicting daily temperature in bandung. In: Proc Int Conf Ind Eng Oper Manag, pp 406–412

Zhang Z (2016) A gentle introduction to artificial neural networks. Ann Transl Med. https://doi.org/10.21037/atm.2016.06.20

Zhang X, Tan SC, Li G (2014) Development of an ambient air temperature prediction model. Energy Build 73:166–170. https://doi.org/10.1016/j.enbuild.2014.01.006

Zhang X, Zhang Q, Zhang G, Nie Z, Gui Z, Que H (2018) A novel hybrid data-driven model for daily land surface temperature predicting using long short-term memory neural network based on ensemble empirical mode decomposition. Int J Environ Res Public Health. https://doi.org/10.3390/ijerph15051032

Zhou Y, Huang Y, Pang J, Wang K (2019) Remaining useful life prediction for supercapacitor based on long short-term memory neural network. J Power Sources 440:227149. https://doi.org/10.1016/j.jpowsour.2019.227149